Claims

1. Method for controlling (CTRL, VDEC) the insertion of additional fields or frames into a first format (24p) picture sequence having a frame frequency of for example essentially 24Hz in order to construct therefrom a second format (25fps) picture sequence the frame frequency of which is constant and is greater than that of the first format picture sequence, e.g. 50Hz, said method including the steps:
   - determining (CTRL, VDEC, ADEC) locations of fields or frames in said first format picture sequence at which locations the insertion of a corresponding additional field or frame causes a minimum visible motion judder (MJT) in said second format picture sequence;
   - inserting (CTRL, VDEC) in said first format picture sequence a field or a frame at some of said locations at non-regular field or frame insertion distances (FRD) such that in total the average distance between any adjacent frames corresponds to that of said second format picture sequence;
   - presenting said first format picture sequence together with said non-regularly inserted fields and/or frames in the format of said second format picture sequence,
   **characterised** in that said field or frame insertion locations in said first format picture sequence are controlled such that, in order to gain perceived lip-sync, in said second format picture sequence the maximum picture content delay caused by the insertion irregularity is kept smaller than average in case a slowly moving or static scene and speech in the audio information assigned to said first format picture sequence are detected.

2. Apparatus for controlling (CTRL, VDEC) the insertion of additional fields or frames into a first format (24p) picture sequence in order to construct therefrom a second

format (25fps) picture sequence the frame frequency of for example essentially 24Hz which is constant and is greater than that of the first format picture sequence, e.g. 50Hz, said apparatus including means (CTRL, VDEC, ADEC) that are adapted
for determining locations of fields or frames in said first format picture sequence at which locations the insertion of a corresponding additional field or frame causes a minimum visible motion judder (MJT) in said second format picture sequence,
and for inserting in said first format picture sequence a field or a frame at some of said locations at non-regular field or frame insertion distances (FRD) such that in total the average distance between any adjacent frames corresponds to that of said second format picture sequence, and for presenting said first format picture sequence together with said non-regularly inserted fields and/or frames in the format of said second format picture sequence,
**characterised** in that said field or frame insertion locations in said first format picture sequence are controlled by said means such that, in order to gain perceived lip-sync, in said second format picture sequence the maximum picture content delay caused by the insertion irregularity is kept smaller than average in case a slowly moving or static scene and speech in the audio information assigned to said first format picture sequence are detected.

3. Apparatus according to claim 2, said apparatus being an optical disc player or an optical disc recorder, or a harddisk recorder, e.g. an HDD recorder or a PC, or a settop box, or a TV receiver.

4. Apparatus according to claim 2 or 3, said apparatus being an optical disc player or an optical disc recorder or a

harddisk recorder or a settop box, wherein said apparatus outputs either the original first format (24p) picture sequence or said second format (25fps) picture sequence, which choice is controlled by replay mode information received either automatically from an interface (IF) that is connected to a device including a display device, or is received from a user interface (UI).

5. Method according to claim 1, or apparatus according to one of claims 2 to 4, wherein speech in the audio information assigned to said first format picture sequence is detected by evaluating, in multi-channel audio, whether the centre channel relative to left and right channels shows a bursty energy distribution over time that is significantly different from the energy distribution in the left and right channels.

6. Method according to claim 1 or 5, or apparatus according to one of claims 2 to 5, wherein said first format (24p) picture sequence is stored or recorded on a storage medium (D), e.g. an optical disc or a harddisk, or is broadcast or transferred as a digital TV signal.

7. Method according to one of claims 1, 5 and 6, or apparatus according to one of claims 2 to 6, wherein said field or frame insertion locations in said first format picture sequence are frames or fields that do not contain large moving picture content areas, the motion being determined by evaluating motion vectors.

8. Method according to one of claims 1 and 5 to 7, or apparatus according to one of claims 2 to 7, wherein said field or frame insertion locations in said first format picture sequence are frames or fields at which scene changes or a fade-to-black or a fade-to-white or a fade to any colour occurs.

9. Method according to one of claims 1 and 5 to 8, or apparatus according to one of claims 2 to 8, wherein the inserted fields or frames are motion compensated before being output in said second format picture sequence.

10. Method according to one of claims 1 and 5 to 9, or apparatus according to one of claims 2 to 9, wherein said first format picture sequence is an MPEG-2 picture sequence and wherein said inserting (CTRL, VDEC) of fields or frames in said first format picture sequence is controlled by evaluating flags either for indicating temporal order of fields or for indicating repetition of the first field for display, which flags are conveyed in said first format picture sequence in a user data field for each picture.

11. Method for facilitating at encoder side a decoder-side control of the insertion of additional fields or frames into an MPEG-2 picture sequence having a frame frequency of for example essentially 24Hz in order to construct therefrom a picture sequence the frame frequency of which is greater, e.g. 50Hz, wherein field or frame insertion locations in said picture sequence are to be controlled by conveyed flags such that, in order to gain perceived lip-sync, the maximum picture content delay caused by the insertion irregularity is kept smaller than average in case there is a slowly moving or static scene as well as speech in the audio information assigned to said picture sequence, said method including the step of inserting, for each picture in said picture sequence, in a user data field either flags for indicating temporal order of fields or flags for indicating repetition of the first field for display.